CDF note 10594

Measurement of the b jet Production Cross Section for events with a Z boson with 9.1 fb^{-1} of data.

The CDF Collaboration
URL http://www-cdf.fnal.gov

(Dated: March 23, 2012)

Preliminary results on the b jet production cross section for events with a Z boson (decaying leptonically) are presented. Measurements are performed using the full CDF dataset. Events are selected by requiring two high p_T electrons or muons and high p_T central jets, among which jets are further required to be b-tagged. The results are compared with predictions from ALPGEN and NLO QCD calculations.

I. INTRODUCTION

Z production in association with b jets is an interesting process not only as it provides a test of pQCD predictions but also because it is a fundamental background for new physics searches. Even more, since theoretical predictions suffer from large uncertainties, dedicated measurements of these processes are therefore crucial.
This report presents a measurement of the ratio of the integrated Z+b-jet cross section to the inclusive Z production and to the integrated Z+jet production cross section. In this study, Z boson candidates are selected via the identification of two high $\mathrm{p_T}$ leptons, either electrons or muons. The invariant mass of the two leptons is required to be within the range [66, 116] GeV/c^2. Jets are reconstructed using the MidPoint algorithm with R=0.7, and the measurement is limited to jets with transverse momentum above 20 GeV/c and rapidity in the range $|Y^{jet}| \leq 1.5$. The measurements are unfolded back to hadron level using the same kinematic requirements for jets and the Z invariant mass.

This analysis which was previously done at CDF using 2 fb^{-1} of data [1] was limited by a statistical uncertainty of $\sim$ 16%. For this new measurement we are using a data sample of approximately 9.1 fb^{-1}. Though the event selection is based on that defined to measure the Z+jets cross section [2] and [3], we have improved the lepton ID definition by implementing an Artificial Neural Network (ANN). The ANN increase the Z boson acceptance by almost 30%. With respect to previous results, the total uncertainties have been reduced by $\sim$50%.

II. MONTE CARLO SIMULATION

Monte Carlo simulated samples are used to model the $Z \rightarrow l^+l^-+$ b-jets signal reconstruction, to estimate background contributions, to evaluate the acceptance and to build the secondary vertex mass templates. These templates are used to discriminate the different jet flavors in the tagged sample (details in the Section V).
The MC samples are described below.

- **Alpgen Z + jets MC** ALPGEN v2.10' [4] interfaced to PYTHIA v6.325 tune BW with CTEQ5L PDFs [5] is used to simulate $Z \rightarrow l^+l^-+$ jets signal events. Alpgen samples are combined using the "MLM matching"scheme and weighted to reproduce the luminosity profile of the data sample.

- **$\mathbf{Z \rightarrow l^+l^-}$ Inclusive MC** PYTHIA Tune A MC samples of $Z \rightarrow l^+l^-$ inclusive covering different data periods are used to evaluate the Z acceptance.

- **Monte Carlo samples for the background estimation** Several sources of background from physics processes are considered. The main contribution to the Z+b-jet candidate sample comes from $t\bar{t}$ events. Other considered processes are diboson (ZZ, ZW, WW) production and $\mathrm{Z \rightarrow l^+l^-} + \gamma$ events. All of these are simulated using PYTHIA 6.216.

III. EVENT SELECTION

Events are collected using a three-level trigger system which requires the presence of a high $\mathrm{p_T}$ lepton (electron or muon). From this sample, only those events with a reconstructed primary vertex with z position within 60 cm from the nominal interaction point are selected for the analysis.

A. Z selection

1. $Z \rightarrow \mu^+\mu^-$ selection with an ANN

For the $Z \rightarrow \mu^+\mu^-$ decay channel a new technique based on Artificial Neural Networks (ANNs) is implemented to improve the muon identification efficiency. This new muon identification technique is based on the discrimination between real high p_T muons coming from a $Z \rightarrow \mu^+\mu^-$ decay and two different sources of fake muons. The first category of fakes comes from tracks of charged particles originated within jet fragmentation. This kind of fakes can be distinguished from real muons because they have similar probability to have same charge or opposite charge with respect to another muon identified in the same event, and for this reason this category is defined as *Same Charge* (SC) fakes. The other category of fake muons comes from low p_T charged particles which undergo a decay in flight, and because of a kink in their trajectory are incorrectly reconstructed as high p_T tracks. This category is defined as *Decay in Flight* fakes (DIF) and can be distinguished from real muons for their high impact parameter and poor quality of the tracking fit. Three different samples of muons corresponding to real muons, SC fakes and DIF fakes are selected from data and two different ANNs are trained to distinguish real muons from fakes.

Muons are thus selected if the SC ANN output is higher than 0.875 and the DIF ANN output is higher than 0.725. This values correspond to the best significance for the inclusive $Z \rightarrow \mu^+\mu^-$ reconstruction. Figure 1 shows the $Z \rightarrow \mu^+\mu^-$ inclusive significance as a function of the output of

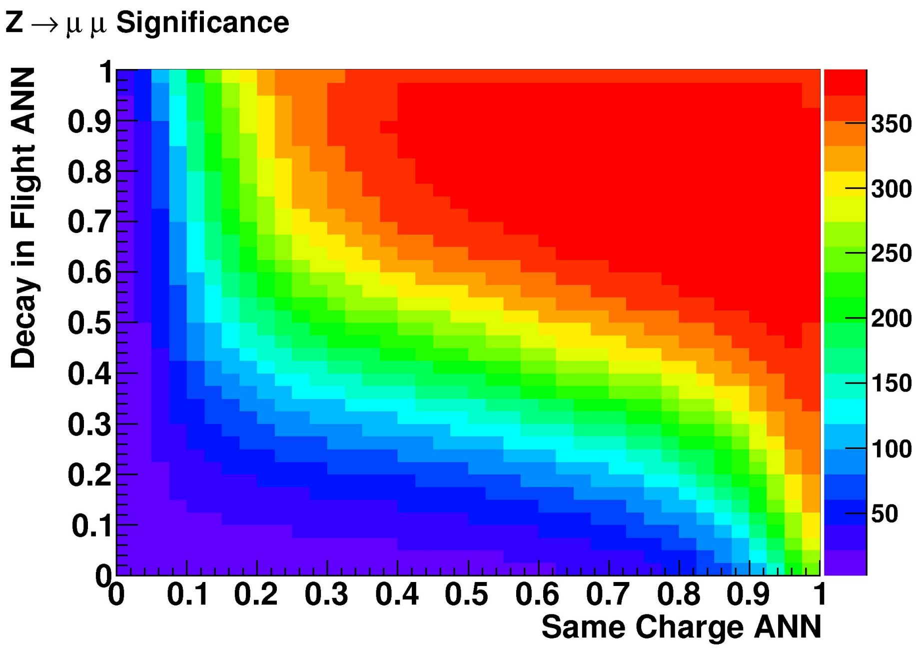

FIG. 1: $Z \rightarrow \mu^+\mu^-$ inclusive significance as a function of the output of the SC and DIF ANNs

the two ANNs.

2. $Z \to e^+e^-$ selection with an ANN

A $Z \to e^+e^-$ boson is identified requiring two reconstructed electrons with $E_T > 20\ GeV$, and with invariant mass in the range $66 \leqslant M_{ee} \leqslant 116\ GeV/c$. One electron has to be central ($|\eta^e| < 1$), while the second electron can be either central or forward with $1.2 < |\eta^e| < 2.8$.
Electrons are identified offline using the ANN technique to discriminate real hight p_T electrons coming from a $Z \to e^+e^-$ decay and fakes from jets misreconstructed as electrons. A sample of fake electrons is obtained from a jet data sample requiring at least one jet ($E_T \geq 20GeV$) and a fakeable object (a jet matched with an electron) whit an invariant mass that falls outside the Z mass peak. Two ANNs are trained, one for the central electrons and another for plug electrons. Central electrons are selected if the Central ANN output is higher than 0.8, while for plug electrons the ANN output should be greater than 0.4. Using thes ANN cuts the electron ID efficiency is 95 % for central electrons and 91% for the plug ones. The gain in the $Z \to e^+e^-$ respect the selection in [2] is $\sim 40\%$.

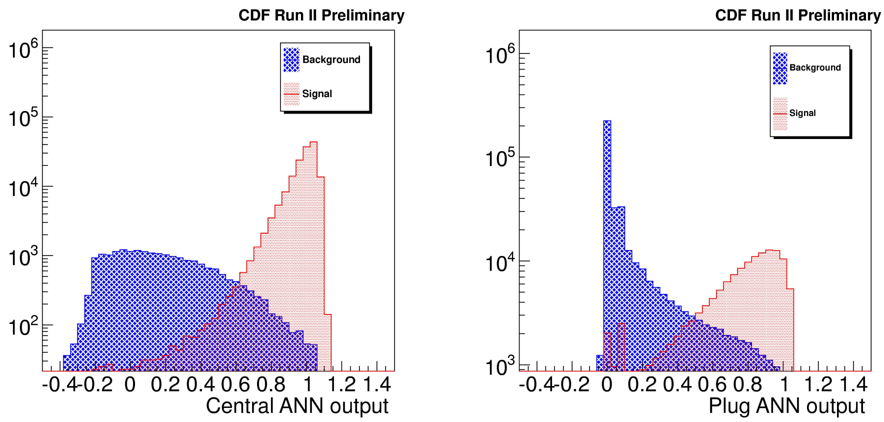

FIG. 2: Output for the two ANNS used to reconstruct the electrons.

B. Jet reconstruction and selection

Jets are identified using the MidPoint algorithm with a cone size of R=0.7 and a merging/splitting fraction set to 0.75. The jets are clustered using calorimeter towers with transverse momentum above 0.1 GeV/c and seed towers of 1 GeV/c. The measured jet transverse momenta are corrected to particle level using the Official CDF Jet Corrections [7]. The measurement is limited to corrected jets with $p_T^{jet} \geq 20\,GeV/c$, and $|Y^{jet}| \leq 1.5$. A minimum distance between the jets and the leptons ($\Delta R_{l-jet} > 0.7$) is also required.

1. b tagged jets

Jets are further required to have a reconstructed secondary vertex within a cone of 0.4 around the jet axis. The secondary vertex is reconstructed using the TIGHT SECVTX algorithm. Jets with a reconstructed secondary vertex are referred to as *tagged* jets. The efficiency to reconstruct a b jet has been measured in data and Monte Carlo and the Monte Carlo efficiency adjusted by a factor 0.96

± 0.05.

C. Observed data events

In the Table I are summarized the events observed for different processes:

	Data	
	muon	electron
N_Z	303 194	540 740
N_{Zjet}	53 941	84 520
$N_{Zbtagjet}$	856	1085

TABLE I: Events observed in data for the electron and muon channels and different processes

IV. THE Z+BJET CROSS SECTION DEFINITION

The b jets cross section for Z+b jets production is defined as:

$$\sigma(Z_{(\to l^+ l^-)} + b\ jets) = \frac{N_{Z+bjet}}{L \cdot A_{Z+bjet}} \tag{1}$$

where the acceptance is the result of a sum over muon/electron categories and over data periods. Its calculation is described in Section VII.
Note that by defining a jet cross section instead of an event one, the measurement does not depend of possible flaws on how our Z+b model handles the physics outside the acceptance.
In order to reduce some systematics (such as luminosity and lepton ID efficiency) the measurement is performed as a ratio of the Z+b-jet cross section with respect to the Z inclusive and the Z+jets cross section.

A. MC Prediction

The requirements, both at generator and reconstructed level, are a Z boson within a mass range 66$\leq M_Z \leq 116\ GeV/c^2$ and b jets with $\mathrm{p_T} \geq 20$ and $|Y| \leq 1.5$.
Summing over all processes in ALPGEN, its prediction (particle level) is found to be:

$$\sigma = 0.294\ pb$$

While the ratio over the Z+jet cross section is:

$$R^{alpgen} = 0.0142$$

B. Measuring the Z+bjet cross section

Since the b-tagged sample is contaminated by charm and light flavor (LF) jets, the number of b tagged jets from Z+b jets production is not equal to the number of tagged jets in data. Therefore the fraction of b-jets, f_b, is determined by fitting a distribution of the secondary vertex mass, M_{SecVtx}, of positively tagged jets to templates for b, c and LF jets. The templates and the contribution from backgrounds, such as $t\bar{t}$ and diboson, are derived from MC.

$$\sigma_{b-jet}(Z+b-jet) = \frac{(n_{data}^{tagged} - n_{data}^{bkg}) \cdot f_b - n_{MC}^{bkg}}{L \cdot A_{Z+bjet} \cdot \epsilon_{tag}^{b}} \tag{2}$$

The vertex mass fit, background subtraction and acceptance calculation are discussed in the following sections.

V. THE B TAGGED SAMPLE COMPOSITION

As explained in previous section, the tagged sample is contaminated by charm and LF jets. In order to estimate the fraction of b jets we perform a fit to the invariant mass of all charged tracks reconstructed at the secondary vertex (M_{SecVtx}).
As one can see in Figure 3 the mass of the tracks forming the secondary vertex is related to the mass of the particles decaying at that point.

Templates are obtained from ALPGEN MC and used to performed a binned maximum likelihood fit to data. The result of this fit is shown in Figure 4.
The fraction of b jets is found to be 0.47 ± 0.04.

We validated the fit by means of pseudo-experiments. We built the SecVtx mass pseudo-data using the same statistics as in data and performed the fit under various b/c/LF scenario. We found no bias and the expected relative error is estimated as 10 % for an scenario similar to the one found in data.

A. Vertex Mass shape systematics

Since templates were obtained from MC, we have determined systematic uncertainties due to possible mismodelings.

1. *Track reconstruction efficiency*

The uncertainty on the template shape due to possible track reconstruction inefficiency is estimated by randomly rejecting 3 % of tracks and recomputing the value of M_{SecVtx} in the Monte Carlo using the tracks inside tagged jets. This represents the main systematics, and it amounts to $\sim 5.5\%$.

2. *Single/Double B/C hadron in a jet*

In ALPGEN the ratio of tagged jets with double to single b (c) quark is 0.23 (0.68). However studies have shown differences with respect to data as large as a factor of 3 (2) times the one obtained

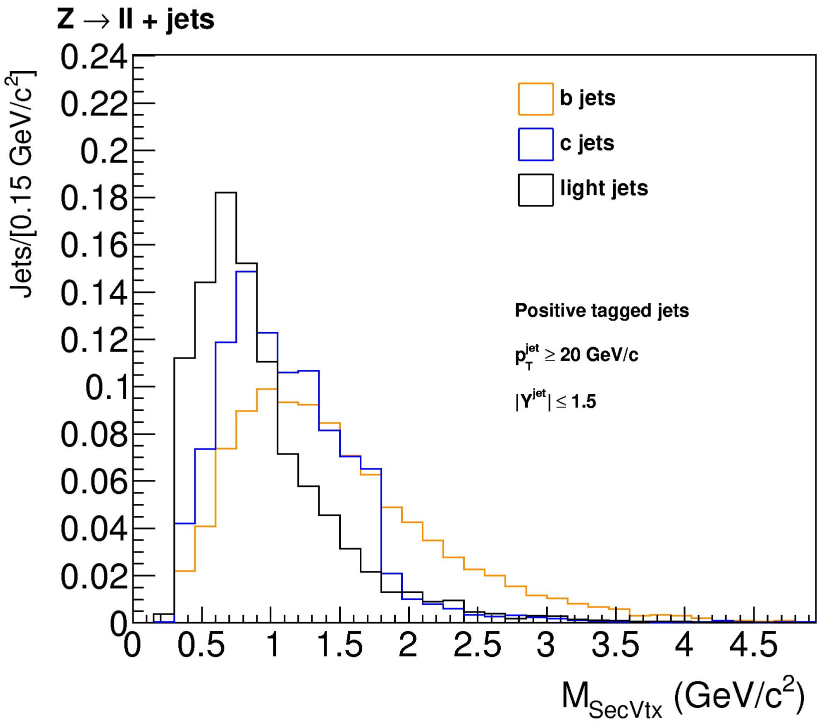

FIG. 3: Secondary vertex mass templates for Tight SECVTX tagged b, c, LF jets

in MC. In order to take into account this effect we built the templates of b and c by varying the ALPGEN prediction of the double heavy flavor cross section by a factor from 0 to 3 (2) i. e. the double fraction varies between 0 to 0.7 for b and between 0 to 1.36 for c, and we estimated the error as half of the difference between the corresponding results.

3. Light template systematic

Because the default light flavor template is built from ALPGEN MC, we have included a systematic calculated from the difference obtained when replacing the nominal light flavor MC template with one constructed using negative tagged jets from data. The contribution to the systematics is ~ 8.7 %.

VI. BACKGROUND SOURCES OF B JETS

Backgrounds were evaluated for the Z inclusive cross section, the Z+jets one and the Z+ b-tagged jet sample. In the case of the Z and Z+jets cross sections, the backgrounds are subtracted from the number of observed data events passing the corresponding selection requirements. In the case of the b-tagged jet sample, sources of b jet background, such as top pair production and diboson, are subtracted from the fitted number of b jets while the background from QCD (fakes) are included directly in the SecVtx Mass fit. This contribution is evaluated from data while others are obtained

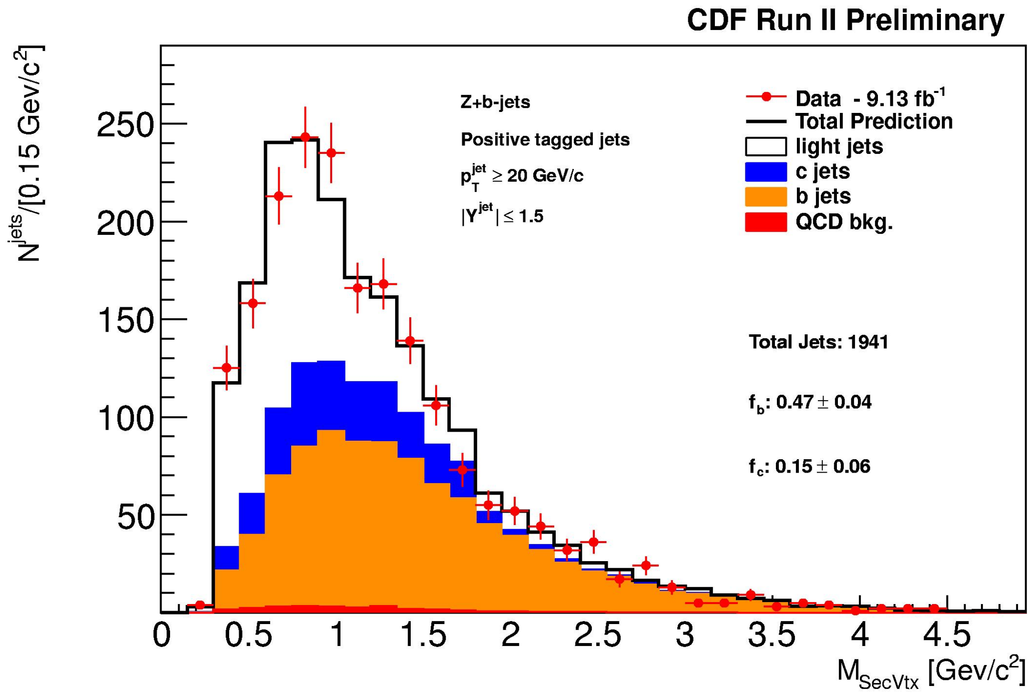

FIG. 4: SecVtx Mass distribution for $Z \to l^+l^-$. It is shown in colors the fitted fraction from the different jet flavors.

from MC.

Expected backgrounds are summarized in Table II and III.

CDF II Preliminary - $L = 9.13 fb^{-1}$

process	bkg estimation
dibosons	30.8 ± 7.5
$Z+\gamma$	18.0 ± 3.7
$t\bar{t}$	133.4 ± 13.3
total	182.2 ± 23.8

TABLE II: The MC background contributions for Z+b tagged jets sample

CDF II Preliminary - $L = 9.13 fb^{-1}$

	Data	Fakes	Bkg Estimation
N_Z	843 934	4987	32 401 ± 6476
N_{Zjet}	138461	1585	5 880 ± 1133
$N_{Zbtagjet}$	1941	28	182±24

TABLE III: Sample composition for the different processes

VII. ACCEPTANCE

The acceptance for each component of this analysis, Z inclusive cross section, Z+ jets and Z+b-jets, was evaluated separately for the muon and electron channel using MC and later combined weighted by each category luminosity. The acceptances are found in the table IV.

A. Acceptance systematics

Several sources of systematic error affect the acceptance determination. The first systematic that is considered is the one due to the Jet Energy Scale uncertainty. The measured energies of jets at CDF are corrected for various effects as already explained in Section III. So this systematic is evaluated adjusting the jet energy corrections within $\pm\sigma$ uncertainties on the nominal jet correction values.

The second source of systematic error in the acceptance comes from the imprecise knowledge of the tagging efficiency. This uncertainty is introduced in the measurement through the error of the b-tag SF (0.96 ± 0.05). Other systematics that affects the acceptances are the uncertainties on the trigger, lepton ID and z vertex efficiency, but these uncertainties are negligible and largely cancelling in the ratios.

CDF II Preliminary - $L = 9.13 fb^{-1}$

Acceptance	electron and muon
A_Z	0.4391 ± 0.0004
A_{Z+jet}	0.527 ± 0.001
$A_{Z+b\_tag\_jet}$	0.140 ± 0.002

TABLE IV: Acceptances for each process (stat errors only).

VIII. SYSTEMATICS

The various sources of systematic errors have been described in the previous sections. All systematics and the corresponding contributions to the uncertainty on the cross section measurement are shown in Table **??**.

The total systematic uncertainty is $\sim$ 11% for the ratio with respect to the Z inclusive cross section, and 13 % for the ratio with respect to Z+jets, both comparable to the statistical uncertainties.

IX. RESULTS

We have obtained preliminary results on the ratios of the b jet cross section with respect to the inclusive Z and Z+jets cross sections. These measurements are defined for events with a Z boson with invariant mass within $66 \leq M_{l^+l^-} \leq 116\ GeV/c^2$ and jets of $p_T \geq 20$ GeV/c and $|Y| \leq 1.5$.

The ratio with respect to the Z inclusive cross section was found to be:

$$\frac{\sigma_{Z+bjet}}{\sigma_Z} = 0.261 \pm 0.023^{stat} \pm 0.029^{syst}\%$$

$L_{int} = 9.13 fb^{-1}$, CDF Run II Preliminary

Systematics	$\frac{\sigma_{Z\_bjet}}{\sigma_Z}$ (%)	$\frac{\sigma_{Z\_bjet}}{\sigma_{Zjet}}$ (%)
Acceptance Systematics		
Jet Energy Scale abs	1.0	2.8
Jet Energy Scale mpi	1.4	6.0
Jet Energy Scale eta	0.6	2.0
b tag efficiency	5.2	
Templates Systematics		
Light Templates - data	8.7	
Double 1b/2b	1.6	
Double 1c/2c	3.3	
Tracks Rec ϵ	5.5	
Others		
Background subtraction	0.9	0.9

and with respect to the Z+jets inclusive cross section :

$$\frac{\sigma_{Z+bjet}}{\sigma_{Zjet}} = 2.08 \pm 0.18^{stat} \pm 0.27^{syst}\%$$

Results have been compared with the ALPGEN prediction and QCD calculations at NLO, as implemented in MCFM [9]. The latter are :

	NLO $Q^2 = m_Z^2 + p_{T,Z}^2$	NLO $Q^2 = < p_{T,jet}^2 >$
$\frac{\sigma(Z\text{+b})}{\sigma(Z)}$	2.3×10^{-3}	2.9×10^{-3}
$\frac{\sigma(Z\text{+b})}{\sigma(Z\text{+jet})}$	1.8×10^{-2}	2.2×10^{-2}

The measured cross section ratios are found to be larger than those from the ALPGEN prediction, by a factor of 1.6, though they are in agreement with MCFM within its large uncertainties. When possible, results have been compared with predictions using different renormalization and factorization scales and found that the measurements preferred those with lower scales.

The differential cross section measurements for Z+b jet as function of jets p_T and rapidity are shown in Figure 5 and 6. They are obtained performing the same measurement described before for each p_T and rapidity bin. These measurements are statistically limited since the statistical uncertainty is $\sim 16\%$.
Measurements have been compared with NLO theory evaluated with MCFM predictions using different renormalization and factorization scales and with different PDF. The prediction is also corrected to take into account non pQCD effects such as hadronization and underlying events. These effects are evaluated with ALPGEN+PYTHIA Tune P2011.
The measurements show a good agreement between data and theory even if they are affected by a large uncertainty.

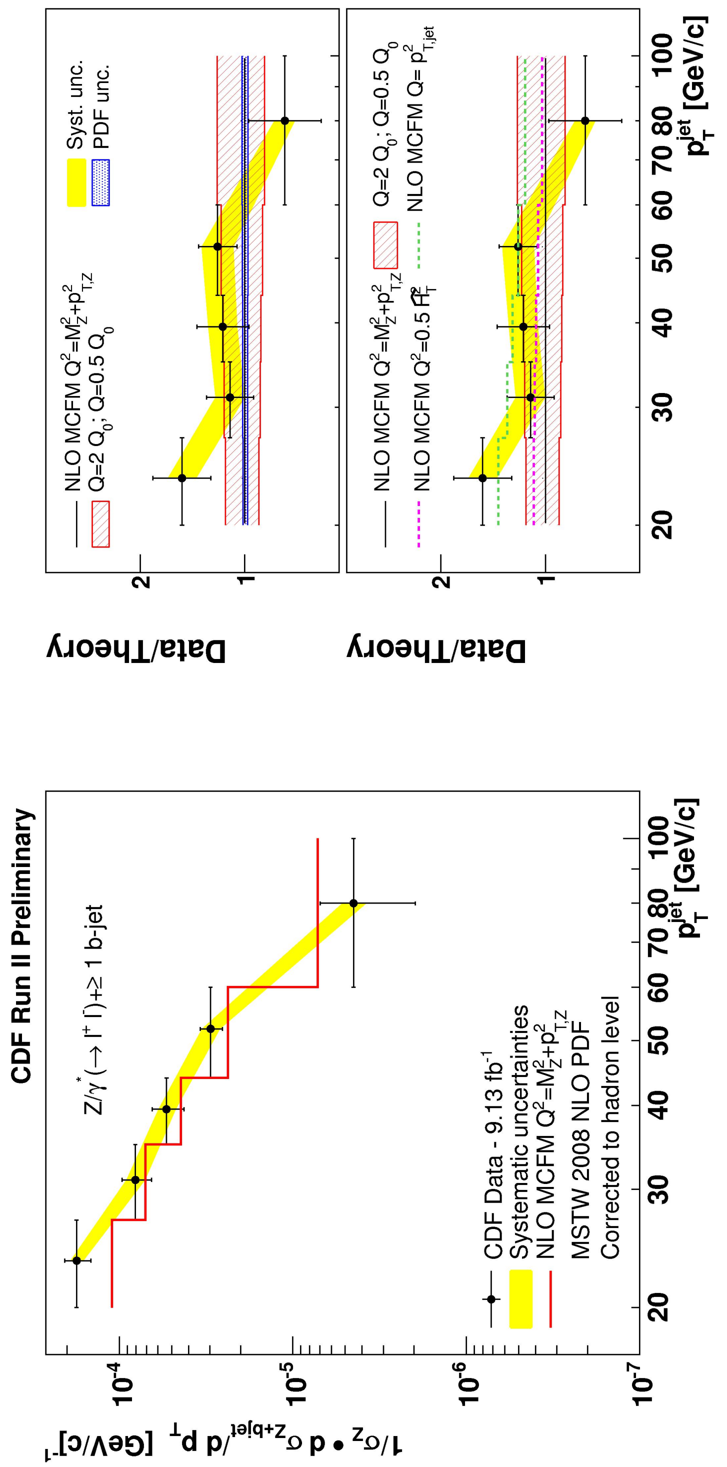

FIG. 5: Z+b jet differential cross section with respect to Z inclusive cross section as function of b jets p_T. The ratio shows the comparison at different renormalization and factorization scales and the impact due to PDF uncertainty and scale uncertainty.

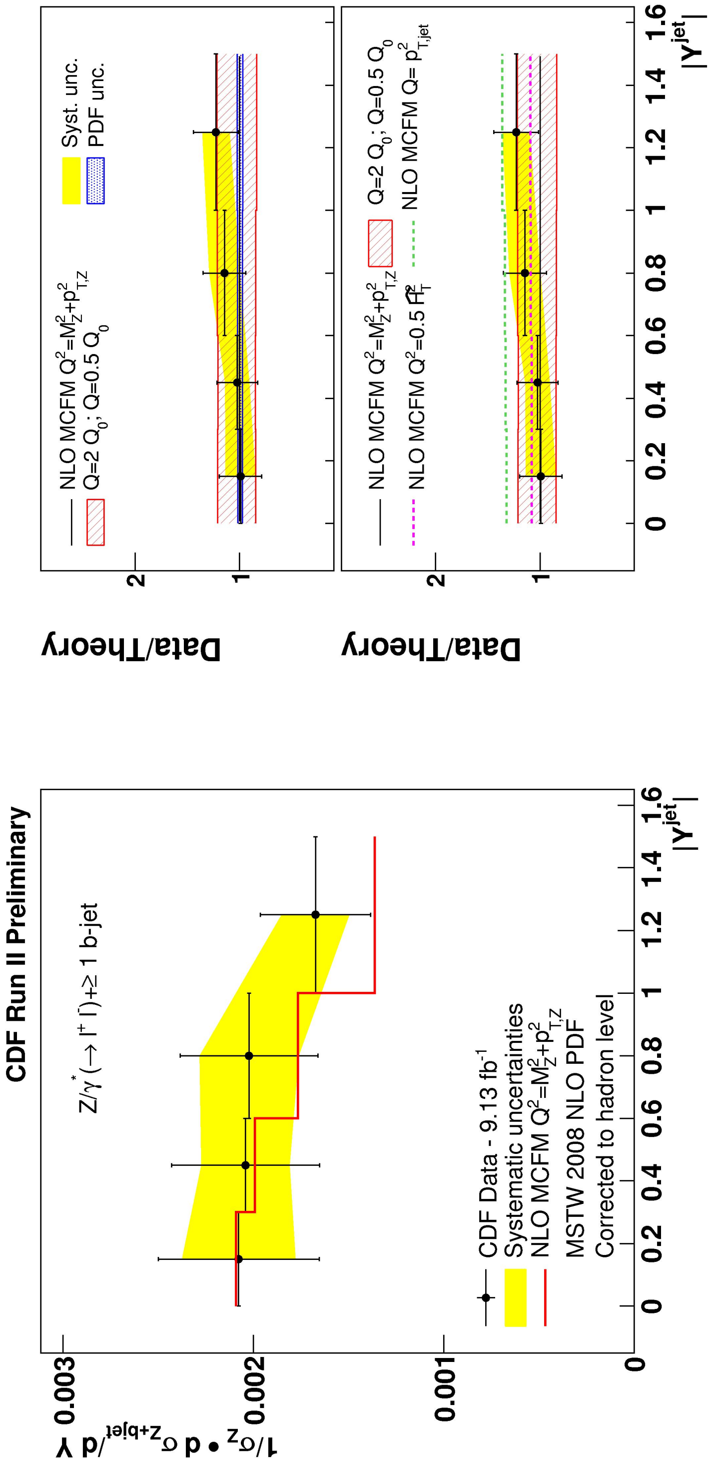

FIG. 6: Z+b jet differential cross section with respect to Z inclusive cross section as function of b jets rapidity.

[1] The CDF Collaboration, Measurement of the b jet cross section for events with a Z boson for 2.0 fb^{-1} Phys. Rev. D **79**, 052008 (2009).

[2] Stefano Camarda, Monica D'Onofrio, Sebastian Grinstein, Mario Martinez, Lorenzo Ortolan, Veronica Sorin Measurement of $Z \to e^+e^-$+jets Production Cross Section with 6 fb^{-1} CDF Note 10394

[3] Stefano Camarda, Monica D'Onofrio, Sebastian Grinstein, Mario Martinez, Lorenzo Ortolan, Veronica Sorin Measurement of $Z \to \mu^+\mu^-$+jets Production Cross Section with 6 fb^{-1} CDF Note 10216

[4] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. Polosa, ALPGEN, a generator for hard multiparton processes in hadronic collisions

[5] H.L. Lai *et al*, Eur. Phys. J. C **12**, 375 (2000).

[6] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. Lett. **94**, 091803 (2005).

[7] A. Bhatti *et al*, Nucl. Instrum. Methods A **566**, 375 (2006).

[8] The CDF Collaboration, Phys. Rev. D **71**, 052003 (2005).

[9] J. Campbell, K. Ellis, MCFM - Monte Carlo for FeMtobarn processes, `http://mcfm.fnal.gov`